 ISBN: 978-989-95618-4-7

Real-Time Incidents Detection in the Highways of the Future

A. Pecharromán, N. Sánchez, J. Torres, J.M. Menéndez

Grupo de Aplicación de Telecomunicaciones Visuales, Universidad Politécnica de Madrid
Avd. Complutense, 30, Madrid - 28040, Spain.
{apg, nsa, jta, jmm}@gatv.ssr.upm.es

Abstract. Due to ever increasing transportation of people and goods, automatic traffic surveillance is becoming a key issue for both providing safety to road users and improving traffic control in an efficient way. In this paper, we propose a new system that, exploiting the capabilities that both computer vision and machine learning offer, is able to detect and track different types of real incidents on a highway. Specifically, it is able to accurately detect not only stopped vehicles, but also drivers and passengers leaving the stopped vehicle, and other pedestrians present in the roadway. Additionally, a theoretical approach for detecting vehicles which may leave the road in an unexpected way is also presented. The system works in real-time and it has been optimized for working outdoor, being thus appropriate for its deployment in a real-world environment like a highway. First experimental results on a dataset created with videos provided by two Spanish highway operators demonstrate the effectiveness of the proposed system and its robustness against noise and low-quality videos.

Keywords: Stopped Vehicle Detection. Passenger and Driver Detection. Pedestrian Detection. Computer Vision. Highway traffic. Adaptive Background Subtraction. Deterministic Rules. Intelligent Transportation Systems.

1 Introduction

Annual traffic crash reports and other statistical information about accidents in traffic give us a lot of insight on how important is for the operators in the traffic control centres the early detection of any abnormal situation, incident or accident, which happened on the stretch of highway under their surveillance. It allows a quick reaction that avoids most of the times worst consequences, helping to save lives in any case.

Bearing this in mind, a novel flexible and robust computer vision-based system for the automatic detection of incidents in highways is proposed in this paper. Functionality is achieved using as input visual information provided by the amount of analog cameras already deployed in this kind of transport infrastructures or, in new deployments currently being carried out, using just only one IP camera. Computer vision, in combination at a higher level of abstraction with the appropriate AI

 ISBN: 978-989-95618-4-7

machine learning approach, has been shown to be nowadays one of the pillars of the automatic detection of incidents [1][20].

There are a lot of commercial systems [2-5] and solutions in the field of the Intelligent Transportation Systems [6][7] for aiding operators in the task of scene understanding, operating on a network of cameras and able to detect automatically some incidents in their field of view. However, the set of functionalities provided by these systems for road safety is still quite limited taking into account all the possibilities that computer vision and machine learning offer. Other disadvantage of these commercial solutions is their limitations as regard as flexibility given that they are closed systems which use own cameras and network interconnections.

Therefore, as an outcome of the conducted research and development, a new application has been developed for supporting operators in Traffic Control Centres with a real-time application for the automatic detection of abnormal incidents on highways, making use of improved state-of-the-art computer vision algorithms for the accurate detection of stopped vehicles and pedestrian present in the roadway, but focusing on giving new functionalities with regards to other existing solutions such as the detection of drivers and passengers leaving the vehicle, which constitutes one of the most risky situations in a highway. Finally, we have simulated some trajectories corresponding to situations such as a car leaving the road path because of an accident or breakdown in order to let the operator be also aware of them.

For such purpose, a method based on pixel history step change and future stability by introducing an artificial delay for the video processing is first applied. The proposed approach is highly based on the method presented in [8], where the detection of moving and stationary foreground objects is able to overcome occlusions in a region of interest by means of a layered representation of the background image. In addition, a robust tracking process is activated for each one of the objects detected in a predefined region to reinforce the high level with the current state of the stopped vehicles in the scene at any time. Reasoning on the current situation is finally made by matching the tracking results against some predefined deterministic rules.

This solution arises in contrast to other stopped vehicle detection solutions [9][10] in which the technique that uses to be applied for background modelling adapts quickly to changes in the environment, being difficult to distinguish between the different types of foreground objects.

The proposed system for the real-time detection of incidents in highways has been shown to work 24 hours a day, 7 days a week. This application has been designed to work accurately outdoors, showing a remarkable robustness against noise due to camera movement, changes in the illumination conditions or a low quality video. Moreover, it has been tested and proved to work well with both real-time videos, acquired with different type of cameras, and compressed videos, as evidenced by the experimental results in Section 5.

To sum up, our paper has been structured according to what has just been described. In Section 2, the overall structure of the system to have a global vision is presented. Some technical challenges are first discussed, such as how the location of target objects in the field of view conditions a lot the detection results. Then, in Section 3, we will go into details about the different modules, first seeing the lowest level of the application, in which background modelling and scene information extraction are included, as well as the accurate identification of the target object. In

 ISBN: 978-989-95618-4-7

Section 4, there is a description of how the system is able to describe the activity being carried out by vehicles and pedestrians in the scene. Finally, some experimental results are presented in Section 5 and some conclusions and future work are extracted in the last section of the paper.

2 General Structure

In this Section, the general structure of the system is presented, which includes the functionality offered by the different modules and the input/output characteristics. The technical challenges the proposed system will have to deal with are first compiled. Then, the paper gets into the problem of the incidents detection.

2.1 Technical Challenges

Dealing with the working conditions is the first technical challenge we have to face, since resolution & quality of the image and orientation of the camera have a big influence on the results we desire to obtain.

For instance, this system does not guarantee a good detection if the target objects are more than 50 meters away from the camera with a normal angle. It is also necessary to avoid reflections caused by the direct projection of sunlight on it. In addition, for the proper maintenance of the extracted background, it is necessary that the camera remains static all the time, i.e., it does not support zoom or pan & tilt movements.

Moreover, the proposed system works in real time, 24 hours a day 7 days a week, so it is expected to have a limited and stable computational cost.

Finally, although several systems have been previously proposed for the detection of incidents in highways [9][10], the functionality they provided is quite limited taking into account the real necessities of the operators. So we are facing the development of a novel system that can provide in a modular way different functionalities such as the following:

- Stopped vehicle detection.
- Drivers & passengers around the vehicle detection.
- Pedestrian in a bounded area of the scene detection.
- Vehicles leaving the road in abnormal circumstances detection

ISBN: 978-989-95618-4-7

2.2 Proposed System.

The application is divided into independent modules. Different units are built to provide several functionalities according to the user preferences: stopped vehicle detection, drivers & passenger in roadway detection, pedestrian detection or road leaving detection. Of course, there are interactions between some of them as shown in the Fig. 1.

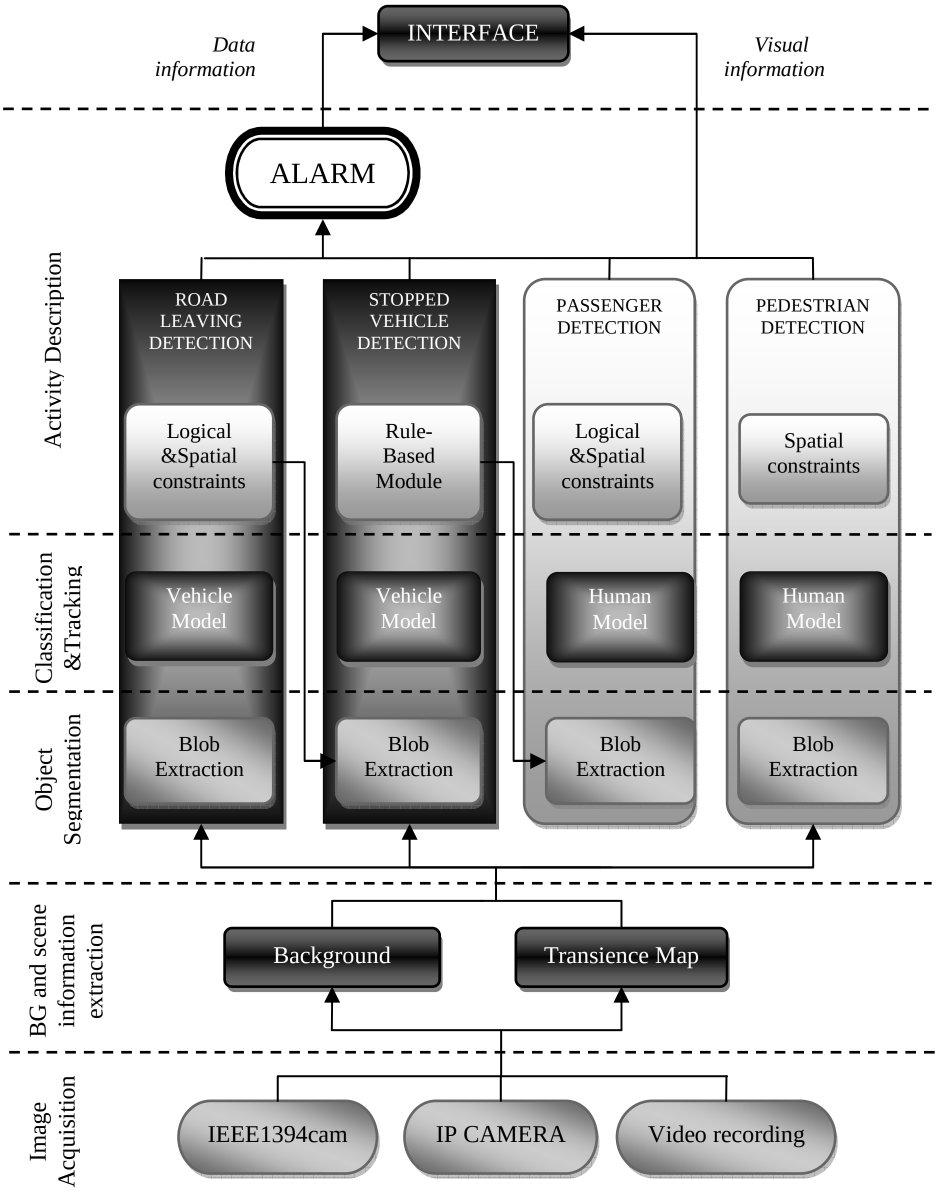

Fig. 1. General Structure of the proposed system

ISBN: 978-989-95618-4-7

First of all, a stream of images is provided by an IP or IEEE1394 Camera located in the proper position, being the section of interest in the highway under surveillance properly covered by its field of view. Background is modelled and both steady and temporary objects are represented in the same image called transience map, thus containing the foreground objects and their transient motion.

An interface was implemented using the Nokia's Qt SDK to let the operator select one or several functionalities the system specifically provides. Once the system has detected the foreground objects in the video sequence, the system will be able to detect whichever stopped vehicles or vehicles out of the road path in the scene.

How the system specifically accomplishes the detection of drivers, passengers and pedestrians in the roadway is a novel methodology susceptible of being patented so no further information can be provided in these moments.

The proposed system is complemented with a module for alarm generation which will launch an alarm in case of a car leaving the road will be detected, if a car stopped in the road has been there for more than five seconds or in case any passenger or pedestrian will be identified in the scene at any moment.

The system will be described in the following Section from the lowest level of processing, i.e. how the background of the scene is modelled for being able to identify subsequently foreground objects, until a description of the activity being carried out is provided.

3 Low-level processing

3.1 Background Modelling and Scene Information Extraction

For the acquisition of the video signal, any IP camera or an IEEE1394 one could be used. Although, as stated in the results section, the system is able to process frames in real-time but also works fine with any kind of compressed video sequences.

After obtaining the images, a map is modelled on the basis of Fujiyoshi et al. [8][14]. After an initialization stage when the background is obtained, each image pixel is classified to represent either background (BG), part of a moving object (TR) or part of a stationary object (ST) according to the pixel history step change value, T, calculated for each image pixel with the following formula:

$$T = \max\{I(t) - I(t-j), j \in [1, N_P]\} \qquad (\mathbf{1})$$

where, I(t) is the intensity value of a given pixel in the currently processed frame, I(t-j) is the value of the corresponding pixel j captured frames before and N_P refers to the number of past frames used to calculate the step change value; and the future stability parameter, S.

 ISBN: 978-989-95618-4-7

The stability value for the pixels of the currently processed frame is calculated with the formula:

$$S = \frac{N_F \sum_{j=0}^{N_F} I(t+j)^2 - \left(\sum_{j=0}^{N_F} I(t+j) \right)^2}{N_F (N_F - 1)} \qquad (\mathbf{2})$$

where, I(t+j) refers to the intensity value of a given pixel j captured frames far in the future and N_F refers to the number of captured future frames used to calculate the future stability value.

The outcome of classifying each image pixel according to T and S is a representation of every pixel in the image as one of the possible states 'BG', 'TR' or 'ST'. It is called transience map.

After this analysis, information from the scene in terms of the number of foreground regions present can be extracted at the object level.

In parallel, background is modelled taking into account the most repeated value for each pixel along a first set of frames used for initialization. Grey scale is used such as in [11] and [12], as using colour would only increase the computational cost, being one of the key challenges in our system.

As it has been previously remarked this system is designed to work 24 hours a day, so it is expected to have also good results even when it is exposed to small changes in the environment, such as lighting and weather conditions.

Therefore an adaptive background like the one proposed in [13] is required. For accomplishing such task, a method which combines the running average technique with a feature selection step in the update process is used. The last step is necessary in order to avoid that foreground objects are included in the background representation. Thus, using the transience map Mt as a selectivity ground, only the image pixel which are in BG state in the Mt are updated to absorb the intensity of the currently processed video frame. The updating of the background image with the selectivity is done according to the formula:

$$I_{BG}(t) = \begin{cases} I_{BG}(t-1) * (1-\alpha) + \alpha * I(t), if\, M_t(t) = BG \\ I_{BG}(t-1), if\; M_t(t) = ST\, o TR \end{cases} \qquad (\mathbf{3})$$

where $I_{BG}(t)$ is the updated background image intensity of a given pixel location at time t, α is the forgetting factor of the background update and $M_t(t)$ is the state of a given pixel location in the transience map at time t.

There are others valid techniques like [15] (as will be mentioned in Section 5) that offer good motion detection results but without an indication of the transient movement of objects, requiring further processing first to discard those regions that not correspond to real objects of interest due for instance to changes in illumination; and secondly to analyse if any object has really become stationary.

 ISBN: 978-989-95618-4-7

After modelling the background the next step is the segmentation of objects in the foreground, so that we can identify objects of interest in the scene for analyzing their activity in next stages.

3.2 Target Object Identification

Object segmentation in the scene is done in different ways depending on which kind of object we want to detect each time.

First of all, in the case of stopped vehicle identification, classification of the stopped vehicle is directly made on the transience map. In it, those blobs labelled as 'ST' (stationary) will be passed through a filter to determine if its size exceeds a minimum number of pixels, and then they are classified as candidates for stopped vehicle. The procedure to determine if a blob corresponds to a vehicle that has stopped in the predefined area is presented in the following section.

For what concerns to people detection, the system is able to detect on one hand, passengers or drivers detected in the area around the stopped vehicle and, on the other hand, pedestrians detected in the area of interest. This part of the paper is under evaluation for being patented so no many details can be provided at this moment.

To segment and accurately classify vehicles that may have left the road, a similar procedure needs to be followed. Having set the area of interest in the calibration stage, the movement detected outside the defined mask will be marked as a vehicle leaving the road in an unexpected way.

After the classification stage, every blob is tracked thanks to its centroid and the distance between it and the centroid of the nearest blob in the next frame. This way, we can do a histogram of the number of times that the tracked blob has been detected as a vehicle or person. Finally, the blob will be treated according to the type that is more repeated in the histogram.

Fig. 2 shows the intermediate results related to the analysis of motion in the sequence.

Fig. 2. Stopped vehicle and driver detection in a Spanish highway. From left to right, up to down, different results can be visualised: final result offered to the operator in the control centre of the highway, transience map, modelled background and pedestrian detection.

4 Activity Description

In an application with so many possible detections like the one proposed in this article, it is necessary establishing an alarm system able to warn the road operator that an incident may have happened. This way, it becomes necessary to know every time what is happening in the road as regard as the set of possible traffic incidents described, i.e., describe the activity in the track.

The activities that the proposed system is able to describe have already been mentioned, but next we are going to describe in more depth how part of this application works as a rule-based system.

The rules followed by the stopped vehicle detector are determined by the number and types of blobs that are extracted at each time in the transience map. Eight possible cases for the system are described according to the pseudo-code presented below.

 ISBN: 978-989-95618-4-7

To understand such code, it is necessary to clarify that *numST* represents the number of blobs labelled as 'ST' in the transience map, *numMX* the number of mixed blobs (according to a percentage in number of pixels established for deciding in a blob is in movement or stationary); and *numTR* represents the number of blobs labelled as 'TR'. Finally, *previousDetection* is a boolean variable that indicates whether in the previous frame there has been detection (then set to true) or not (set to false in that case).

To sum up the following code, as soon as a stationary blob which fulfils the specified size constraints is detected, it is considered as a detection (in the subsequent frame *previousDetection* will be set to true) and marked as a stopped vehicle in the visualization window. Then, depending on the number of 'ST' blobs and if the blob under analysis is bigger or smaller than in the previous frame, we will consider the fact of increasing or decreasing the number of detections bearing in mind possible occlusions.

```
if(!previousDetection){
    CASE 0; //Nothing happens.
    if(numST=!0) CASE 1; //New detection
}else if(previousDetection){
    if(numST==0){
       if(numMX==0){
           if(numTR==0) CASE 2; //State impossible to reach
       Delete.
           else CASE 3; //Delete non occluded
       }else{
           if(numTR==0) CASE 4; //Delete moving
           else CASE 5; //Delete moving before non occluded
       }
    }else{
       if(numMX==0){
           if(numTR==0){
               if(numST>numSTprevious) CASE 6; //New detection
           }else{
               if(numST>numSTprevious){
                   CASE 7 Detected; //New Detection
               }else if(numST<numSTprevious){
                   CASE 7 Deleted; //Delete non occluded
               }
           }
       }else{
           if(numST<numSTprevious){
               CASE 8 Deleted; //Delete moving before non
           occluded
           }else if(numST>numSTprevious){
               CASE 8 Detected; //New Detection.
           }
       }
    }
}
```

If a stopped car is detected in the same position for more than five seconds, an alarm is sent to the road operator in the control centre, and a reminder is also sent if it has been stopped for more than a minute.

 ISBN: 978-989-95618-4-7

For vehicles leaving the road detection, the application also behaves like a rule-based system. There are four possible and simple cases representing the state of the scene:

- First case, there is no object out of the road.
- Second, there is a detection of a vehicle leaving the road (an alarm would be sent to the road operator).
- Third, a vehicle previously detected out of the road is detected moving to the inside of the track.
- Forth, this same vehicle is detected stopped on the track (an alarm would also be sent).

Pedestrian and passenger & driver detection are associated to some spatial constraints. This implies that depending on the detection area the person is in, it would be described as a pedestrian or as a passenger/driver.

5 Results

First, we should state that the computer used to run the application is Intel(R) Core(TM) 2 Quad @2,40GHz 1,99GB of RAM, Windows XP. Besides, our method was implemented using the C++ language and the OpenCV library due to its efficient. The interface was implemented with the Qt 4.2.0 Nokia library making it more intuitive for the operator who will make use of it in their daily activity.

The acquisition of the video during the testing phase in the real scenario is being performed using mainly an AXIS IP camera. However, as already mentioned, the system has been successfully tested with different highly compressed videos in ".avi" format of 704x576 pixels resolution, which allows us to state that this system achieves a good performance even with low quality videos.

The best assembly for the camera with regard to the road it should be situated at a height of 8 to 10 meters, with a tilt angle relative to the horizon equal to 5 and using a lens that, depending on the final height varies between 4.8 and 8mm. That is, this configuration will lead to a balance between the area covered by the field of view of the camera and the resolution that the objects in the scene are represented with. According to an initial analysis, the scope of the algorithm usually ranges from the first 8 to 50 meters from the axis of the camera, to a range of 50 to 100m depending on its positioning.

Video samples were captured in different scenarios (i.e. real highways in Spain), always respecting the maximum distance between the camera and the vehicle to be detected. The application has been tested for periods of time from a few minutes to several hours, having obtained very satisfying results.

In the following table, it can be seen the results for the four kinds of detection after testing the algorithm with 14 videos of different lengths (from 3 to 15 minutes) with different characteristics (including pedestrians walking through the scene, passengers leaving a stopped vehicle, different kind of vehicle stopping in the road and so on), for different real highways in optimum visibility conditions, which have been obtained thanks to the collaboration with [18] and [19].

 ISBN: 978-989-95618-4-7

The percentage present is the average of the percentages in all the videos, not the percentage extracted from the values offered next to it. If we consider the non-detection as the most dangerous problem, it is obvious that it is a very low percentage compared to percentage for the good detection, which is quite high. Despite a 37,04% average of false alarms of stopped vehicles may seem high, those false alarms do not last more than one frame, so in these tests the system has never sent an alarm to the controller because of a stopped vehicle that does not exist.

Moreover, in the table it is shown how our application has been compared to another detection of foreground objects system based on Mixture of Gaussians proposed by Stauffer et al. [15], which generally offers worse results, because despite it usually presents a good motion detection performance, there are too many false alarms.

Table 1. Experiments on the system

		Stopped Vehicle	Passenger	Pedestrian	Off-Road
		% average	% average	% average	% average
Layered	Hit	51,85	85,71	100	16,67
	False Alarm	37,04	0	0	83,33
	Miss	11,11	14,29	0	0
MoG	Hit	19,94	64,29	55,61	50
	False Alarm	70,06	0	0	50
	Miss	10	35,71	44,39	0

The proposed system has high robustness against noise and supports low-quality videos (which have been compressed or not). Obviously a better detection will be expected with the best possible conditions.

Next there are presented some examples of several detections in different scenes. For instance, in Fig. 3 there is a stopped vehicle detected; in Fig. 4, there is a frame in which a passenger around the stopped vehicle is detected and another frame in which a pedestrian is detected out of the area for passengers.

ISBN: 978-989-95618-4-7

Fig. 3. Stopped Vehicle Detection at the entrance of a tunnel of M-12 tool road in Madrid (Spain)

Fig. 4. Passenger and Pedestrian Detection in one of the sections of the Aumancha highway in Toledo (Spain)

ISBN: 978-989-95618-4-7

6 Conclusions

In this paper, a new system for the real-time detection of traffic incidents on the road has been presented, which offers good detection results and which is expected to be improve with further testing and under more complex scenarios.

The system is capable of processing in real-time images from any kind of IP camera or from a compressed video sequence stored in disc, and of detecting four types of incidents at the same time: stopped vehicles, driver & passenger around the vehicle, pedestrian in a bounded area of the track and vehicles leaving the road in an unexpected way. It also always knows the situation of the scene in terms of people or vehicles. Although there is an alarm system for more serious incident, the presence of a road operator becomes necessary.

The modular structure of the program allows future research, since new functionalities can be added to the system (E.g. vehicles driving in opposite detection), making it an even more flexible application. A module for animal detection or classification between different types of vehicles (trucks, cars, motorbikes…) could be also useful in future researches.

7 Acknowledgements

The information / results presented in this article are the result of research conducted within the OASIS project, funded by the Centre for Development of Industrial Technology (CDTI) within the CENIT program. These results belong exclusively to the companies promoting the project and who are part of the Economic Interest Grouping (AIE), this is: OASIS CENIT, A.I.E: IRIDIUM, OHLc, ABERTIS, SICE, Indra, Dragados, OHL, GEOCISA, GMV, ASFALTOS AUGUSTA, HIDROFERSA , EIPSA, PyG, CPS, AEC, Torre de Comares.

8 References

1. VSAM, Section I, "Video surveillance and monitoring", Proc. DARPA Image Understanding Workshop, vol.1, pp.1-140, (1998).
2. Agent Vi, URL: http://www.agentvi.com/.
3. Autoscope, URL: http://www.autoscope.com/.
4. Citylog, URL: http://www.citilog.com/.
5. Ioimage, URL: http://www.ioimage.com/.
6. Hu W., Tan T., Wang L., Maybank S.: A Survey on Visual Surveillance of Object Motion and Behaviours. In: IEEE Transactions on Systems, Man and Cybernetics, vol. 34, pp. 334 – 352 (2004).
7. Haering N., Venetianer P. L., Lipton A.: The Evolution of Video Surveillance: an overview. Machine Vision and Applications, vol. 19, issue 5-6, pp. 279–290 (2008).
8. Fujiyoshi H. & Kanade T.: Layered Detection for Multiple Overlapping Objects. In: 16th International Conference on Pattern Recognition, vol. 4, pp. 156-161, 11-15 August, Quebec, Canada (2002)

 ISBN: 978-989-95618-4-7

9. A. Bevilacqua and S. Vaccari, "Real Time Detection of Stopped Vehicles in Traffic Scenes", IEEE International Conference on Advanced Video and Signal-Based Surveillance, pp. 266 – 270, London, UK (2007).
10. Hu W., Tan T., Wang L., Maybank S.: A Survey on Visual Surveillance of Object Motion and Behaviours. IEEE Transactions on Systems, Man and Cybernetics, vol. 34, pp. 334 – 352, (2004).
11. McKenna S., Jabri S., Duric Z., Rosenfeld A., Wechsler H.: Tracking Groups of People, Computer Vision and Image Understanding, vol. 80, pp. 42-56 (2000).
12. Haritaoglu, Harwood, D., Davis. L.: W4: Who, when, where, what: A realtime system for detecting and tracking people. In: IEEE International Conference on Automatic Face and Gesture Recognition, pp. 222–227, (1998).
13. Monteiro G., Marcos J., Ribeiro M. Batista J. Robust Segmentation for Outdoor Traffic Surveillance. In: 15th IEEE International Conference on Image Processing, 12-15 October, San Diego, USA, pp. 2652-2655 (2008).
14. Kujanperä V. Highway video surveillance – development of an automatic detection system for abnormal incidents employing computer vision. University of Oulu, Department of Electrical and Information Engineering. Master´s Thesis (2010).
15. Stauffer, C.; Grimson, W.E.L.: Adaptive background mixture models for real-time tracking. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol.2, pp.246-252 (1999)
16. Lipton AJ, Fujiyoshi H, Patil RS.: Moving target classification and tracking from real-time video. In: Proceedings of the IEEE workshop on image understanding, pp. 8-14 (1998).
17. Leo M., Spagnolo P., Attolico G. & Distante A.: Shape Based People Detection for Visual Surveillance Systems. In: Audio- and Video-Based Biometric Person Authentication, Springer Berlin, Heidelberg, Germany, pp. 285-293 (2003).
18. OHL, URL: http://www.ohl.es/.
19. Aumancha, URL: http://www.aumancha.com/
20. Sánchez N. and Menéndez J.M.: Video analysis architecture for enhancing Pedestrian and driver safety in public environments. In: 10th International Workshop on Image Analysis for Multimedia Interactive Services, pp.288, London (2009)